May 11, 2001

U.S. CMS
Software and Computing Project

Progress Report
2nd Quarter FY2001

I. Summary

This document is the second U.S. CMS Software and Computing progress report on project status and performance to the funding agencies and to project oversight. It covers the second quarter of FY2001, i.e. the months January to March 2001. The project is progressing well, both in technical achievements and in management of resources.

The project receives significant funding in FY2001 to support computing facilities for U.S. CMS users, to perform computing prototyping, R&D, and core application software development. Part of the funding allocated for FY2001 was received in February and the remaining DOE funds were received in April. A loan from the U.S. CMS detector project is used to cover part of the expenditure to support on-project engineers. This is being implemented by reporting effort to the U.S. CMS Construction Project, unloading those costs from the Software and Computing Project. The funding requested from NSF for FY2001 amounts to $1500k, however only $320k have been received yet, and another $80k were advised. This poses severe problems and forced us to stage the deployment of the prototype Tier-2 equipment and the related R&D program. The funding situation is summarized in Tables 3 and 4 in the funding section.

The U.S. CMS Software and Computing project strengthens its support of the U.S. CMS user community in extending the emerging user facilities. A plan is under development to upgrade the Tier-1 User Facility by providing interactive services and user access to large data sets on storage servers with a multi-Terabyte user object database federation. The R&D related to this proposal was performed during this quarter, including evaluation of storage servers and RAID array performance, Linux cluster configuration etc.

Valuable experience was gained using the CMS fully functional prototype software for production and reconstruction of simulated events on the Tier-1 Linux cluster and storage servers. The production efforts to produce 1.4 Million simulated and fully reconstructed events has important R&D aspects, besides delivering event samples requested by the Physics Reconstruction and Selection study groups. It provides real-world experience with using a worldwide distributed object database (note that the “database selection” milestone is scheduled on the timescale of a year). In addition, important lessons about feasibility of system configurations, system throughput, fault tolerance and possible bottlenecks are being learned by setting up the object reconstruction of simulated events and pile-up simulation. System simulations based on the MONARC simulation toolkit support these studies. In parallel with production running and operations, work on developing a distributed production system is progressing.

Due to lack of funding, deployment of the hardware of the prototype Tier-2 regional center was staged. The first phase was finished at Caltech and UC San Diego. The system was commissioned and the computing facilities were configured ready for CMSIM simulations, including software installation of the CMS environment.

Work proceeded on developing a project plan for Tier-2 and Grid computing related R&D. We expect to finalize the WBS and schedule soon, and the project will then sign SOWs with the Tier-2 institutions to agree on efforts to be delivered, on schedules, funding and reporting. Improving the funding situation would help to plan and manage this important effort. The site for the second prototype Tier-2 center was selected to be University of Florida. CMS is involved in the U.S.

Grid Projects PPDG, GriPhyN and iVDGL. CMS deliverables and requirements are part of the project proposals that were submitted to DOE (for PPDG) and NSF (for iVDGL).

We also report significant progress in the Core Application Software sub-project, where the effort concentrated on developing the CMS Software Architecture and the distributed data management and processing.

In February the final report of the CERN LHC Computing Review was released. The findings and recommendations of the review panel steering committee are very much in line with the U.S. LHC computing plans. The report recommends a prototyping phase until 2004 with significant funding for computing hardware starting this year, and deployment of the computing facilities over three years starting in 2005. M&O and upgrades should be based on a rolling replacement of equipment after 2007, at a constant budget in the order of one-third of the initial investment. The need for software engineering support was recognized and the estimated level of effort is consistent with the U.S. project plan. These issues were brought to the attention of the CERN Resource Review Board (RRB) in April, and will be addressed by an interim MOU between CERN and the funding agencies on software and computing at the end of this year. This should help to secure funding for manpower and resources until the final MOUs for computing will be put into place around 2004.

At CERN the CMS computing and core software efforts have been re-organized and strengthened, in preparation for the LHC research phase. Computing and Core Software (CCS), Physics Reconstruction and Selection (PRS), and Trigger and Data Acquisition (TriDAS) are now organized in a coherent project framework, CPT, with cross-project task forces and a common Technical Board. Tasks, schedules and deliverables for CPT are being defined and will be subject of the iMOU.

CERN has also come up with a schedule for LHC start of operation. First collisions are expected in April 2006, followed by a one-month pilot run at luminosities around $3x10^{32}cm^{-2}s^{-1}$. After a short shutdown a first seven months long physics run will start in August 2006 at expected luminosities of up to $2x10^{33}cm^{-2}s^{-1}$. This will allow accumulating an integrated luminosity of 10 fb^{-1}. The new schedule requires fully functional software and computing environments in spring 2006 and a large part of the computing resources to be operational for the physics run.

II. Technical Status

In this section we report effort and technical achievements. Spent effort is assessed on the basis of regular effort reports from engineers and managers, using a bottom-up approach, where for each WBS item expenditure of labor is accumulated. For the User Facilities subproject this is reconciled with the monthly effort reports from the departments of Fermilab Computing Division. As part of the reporting for the Core Application Software subproject the engineers submit monthly reports for effort tracking.

In the following sections we use the unit of 1 FTE to denote labor of one full-time equivalent working for one year.

The project has made significant technical progress in the User Facilities (WBS 1) and the Core Application Software (WBS 2) sub-projects. In the User Facilities subproject work went on in

several areas, including Monte Carlo production, collaborative tools, software environments, installation and support, documentation, data test beds, and user support.

WBS 1.2 System and User Support

About 0.32 FTE were spent during this quarter on systems and users support. Web server and help desk support required only little effort, as — once setup — central Fermilab/CD groups provide these services. As an ongoing responsibility the CMS code libraries and software development environment was maintained and kept up to date. This included updating of the SCRAM configuration system and installation of new compiler versions, needed to be compatible with new software versions.

Concerning training of support staff, four U.S. CMS computing professionals attended C++ courses.

The prototype Tier-2 center at Caltech and UCSD was supported by providing general consulting and help to install the tools and environment for CMS production.

WBS 1.3 Operations and Infrastructure

Apart from general operations and infrastructure support planning for re-siting the CMS cluster in the computer room required some effort, which totaled to about 0.02 FTE.

WBS 1.4 Tier 2 Regional Center

In this quarter the installation and commissioning of the first-phase hardware and CMS software systems was finished both at UC San Diego and at Caltech.

Effort was spent in hardware procurement and installation, software maintenance and configuration, and system testing. Two high performance RAID arrays were installed and configured, one at each site. This required more time than expected because incompatibilities were found in the drivers used by the CMS default version of Linux and the new high bandwidth SCSI controllers used to access the RAID arrays. A low-cost, medium-performance ATA-100 EIDE based RAID disk server was installed and configured at Caltech for evaluation as a possible high capacity storage solution.

The PBS batch queuing system was enabled for use in CMS production and the computational nodes were configured for production use. CMS regular software maintenance and upgrades were performed. Several generations of the Fermilab production scripts were installed. The original set was fairly site dependent and required work to port to remote centers, the current version is much more generic and easier to use.

Tests were performed running CMSIM production to commission the system. This varifies the software installation, the port of the production scripts, the queuing system, and the archiving tools. A large-scale stability test was also performed using the GDMP grid tool. 200 GB of Objectivity database files were staged from the CERN mass storage system, transferred to California, and automatically integrated into our local Objectivity database federation.

Despite these successes, there was significant lack of manpower due to that NSF funding still did not materialize. The effort spent was mostly off-project. In total it amounted to 0.20 FTE in this quarter.

WBS 1.5 Networking

The Fermilab onsite network interconnecting the Linux farm cluster nodes was designed.

Network performance was monitored throughout the production activities, using the new Web-based tools. These allow logging performance data and analyzing performance and throughput of individual network ports. The total effort covering this item was 0.02 FTE.

WBS 1.6 Computing and Software R&D

R&D in computing hardware for Tier-1 and Tier-2 centers included design of computing systems for test-beds installation. The infrastructure needed for testing farm nodes was setup. Different RAID-array systems were evaluated.

For the Fermilab production cluster very high throughput and CPU-intensive computing is required in order to perform CMS pile-up simulation. R&D and tests were performed to establish ways to sustain the required I/O-rates of about 40MB/sec to and from the Object database. This task was achieved by setting up a system consisting of pile-up servers and worker nodes. The design and performance was documented, production was routinely run and monitored with project-developed tools.

Concerning data access and distribution, work was continuing on a distributed production system. The production system was also simulated using the MONARC simulation tool kit, which allows comparing relevant performance figures with measurements to assess possible bottlenecks and scalability issues and help to develop the system design further.

In this category we also report the system administration for the development clusters.

The total effort amounted to 0.34 FTE.

WBS 1.7 CMS Detector Construction Phase Computing

This item covers running the CMS computing facilities, system administration and maintenance of the CMS software distribution. Some effort was spent on data storage including software interfaces to the mass storage, and distribution of the large data sets including work on the file replication between Fermilab and U.S. institutions and CERN. The total effort was 0.49 FTE.

WBS 1.8 Support for Fermilab Based Computing

There was very little effort left to cover installation of and support for desktop systems. This caused some problems witch will need to be addressed. The effort spent was 0.01 FTE.

The total FTE effort spent on the User Facilities sub-project is summarized in Table 1. These numbers include on- and off-project contributions. The total project funded effort in this quarter was 1.0 FTE.

WBS item	FY2001Q1	FY2001Q2	Total effort projected FY2001	Total project-funded effort FY2001
WBS 1.1 Tier 1 Regional Center			inactive	
WBS 1.2 User Support	0.26	0.32	1.31	
WBS 1.3 Maintenance and Operations	0.03	0.02	0.63	
WBS 1.4 Tier 2 Regional Centers	0.25	0.20	2.50	
WBS 1.5 Tier 1 Network	0.02	0.02	0.50	
WBS 1.6 Software and Computing R&D	0.44	0.34	2.69	
WBS 1.7 Detector Construction Phase Computin	0.41	0.49	2.00	
WBS 1.8 Support for Fermilab Based Computing	0.03	0.01	0.38	
Total FTEs at Fermilab Regional Center	**1.19**	**1.20**	**7.50**	**4.05**
Total FTEs	**1.44**	**1.40**	**10.00**	**6.55**

Table 1 *Summary table for labor effort spent on the User Facilities subproject.*

WBS 2.1 Software Architecture

Significant technical progress is reported in the CMS software architecture development effort for this quarter. This task involves software architecture documentation and evaluation as well as development in the CMS central framework, CARF, and specialized software frameworks for analysis and detector reconstruction.

The first version of the CMS Architecture Forum for Evaluation (CAFÉ) documentation tool was released for use. These tools will be used to ensure that all sections of the software evaluated by CAFÉ have consistent and complete documentation. Work is in progress on the top-level architectural description document for CARF.

Work is progressing on the understanding of the format for storing the detector description database. This is the database for storing the detector geometry and material information that will be used by a variety of CMS software clients including ORCA, OSCAR, IGUANA, and FAMOS. There has been considerable effort spent investigating the use of XML as the storage format, because of the apparent advantages of aspects like the ability to read and edit XML document using simple text editors. As a result of this R&D it is becoming clear that as the detector description becomes more complete and more complicated these benefits of XML may be diminished. A number of alternatives are being investigated. Tools have been developed to convert the CMS GEANT3-based geometry files to XML in an automated way. A spin-off from this work has been investigations into storing parameter information for the endcap muon charged strip chambers using XML. While this is still an investigative effort at the moment, the tools are being designed to be general enough for use with other detector sub-systems.

The IGUANA project entered the fully functional prototype phase with the first evaluation release of the new analysis and visualization architecture. The new software architecture is built

around a very small kernel, which can load a variety of plug-ins. This allows enormous flexibility to build diverse analysis tools around a central architecture.

The total effort amounted to 0.44 FTE.

WBS 2.2 Interactive Graphics and User Analysis

Technical progress on the sections of the analysis tool development not related to architecture largely stalled throughout the second quarter due to a lack of manpower. The funding for the scheduled second engineer for this sub-project has not yet materialized. In addition, the CERN supported software engineer in charge of software configuration and software configuration tools support and development left the experiment at the beginning of the quarter. Without support for software configuration CMS software including IGUANA can no longer be compiled. As an emergency measure, a large percentage of the sole IGUANA developer had to be temporarily diverted to configuration management. This situation improved in April when a new hire at CERN took over the code management duties. The U.S. CMS code librarian at the Fermilab Tier-1 center also started to work on configuration management and release with the support of off-site configuration and distribution issues in mind. Two developers will help to protect CMS from losing expertise with the loss of a single person.

Work has progressed on event visualization with the CMS Physics Reconstruction and Selection Groups, PRS. At the completion of the Functional Prototype Phase about half of the desirable reconstructed physics objects could be visualized, work is in progress to visualize the other half. There has been considerable effort in support and training, so that physics users will be able to participate in visualization development.

The total effort amounted to 0.06 FTE.

WBS 2.3 Distributed Data Management and Processing

This sub-project has short-term goals for the development of tools to facilitate CMS production today and longer-term goals to make efficient use of a globally distributed grid of computing resources for production and analysis. There is progress to report in both areas. CMS has been attempting to finish an ambitious program of simulated event production that began in the first quarter of 2001. During this production events were simulated at computing centers around the world and for the first time were reconstructed at selected regional centers away from CERN. In light of this experience CMS has evaluated their production tools and started to develop more consistent and robust tools. The first step has been to generalize the tools developed by the CAS engineers at Fermilab in support of the U.S. production efforts, by removing the site dependencies. These have become the basis of a consistent set of production tools that are currently being used at CERN, Fermilab, and the California Tier-2 prototype center. There are commitments from the Italian and UK groups to work towards using a common set of tools. Work is currently in progress on a more advanced set to improve the ease and reproducibility of job specification.

There is also progress toward the longer-term goals in the areas of distributed job submission and distribution of dynamic services. An experimental distributed queuing system, developed primarily by an off-project developer at Caltech, has been implemented for testing and

development at the California Tier-2 prototype. The scheduler aims at improving the tracking of long running jobs, facilitating jobs control by a group of users to improve collaboration, conserving limited network bandwidth, maintaining high availability, and improving fault tolerance to network failures.

Other R&D efforts include evaluation of a distributed server architecture for dynamic services, based on JINI network technology. This flexible architecture could be used to exchange information for use by a variety of clients; distributed job schedulers and smart data movers are just two examples.

The total effort amounted to 0.75 FTE.

WBS 2.4 User Support

The support task involves developer and user support, as well as the development and support of software configuration and management tools. To aid in developer support a CMS technical mailing list was created this quarter at cms-techical-computer@listser.fnal.gov. It is designed to give developers a forum to pose questions. To support users there was an ORCA tutorial held at CERN this quarter. It included a very successful IGUANA tutorial, which walked users through the process of modifying the program to visualize their own objects. The CERN tutorial had 50 people attempting to register for 22 available spots. There are plans to schedule another U.S. tutorial at Fermilab sometime this summer. The support and development of tools for software configuration and management received a boost this quarter with the addition of a portion of one UF engineer's time devoted to this task, as mentioned above. Adding the extra developer is extremely valuable and the use of UF efforts located at Fermilab ensures that U.S. users enjoy the same quality of software configuration and support as users at CERN.

The total effort amounted to 0.75 FTE. This number is higher than average due to the overlap in configuration management support and raining.

The total FTE effort spent on the Core Application Software sub-project is summarized in Table 2. The total project funded effort in this quarter was 2.0 FTE.

WBS item		FY2001Q2	Total effort projected FY2001	Total project-funded effort FY2001
WBS 2.1 Software Architecture		0.44		
WBS 2.2 Interactive Graphics and User Analysis		0.06		
WBS 2.3 Distributed Data Management and Proc		0.75		
WBS 2.4 User Support		0.75		
Total FTEs	**0.00**	**2.00**	**8.50**	**7.75**

Table 2 *Summary table for labor effort spent on the Core Application Software subproject*

III. Financial Status

Table 3 shows the funding status of the U.S. CMS Software and Computing Project. Funds received in FY2000 and funds requested for FY2001 are shown. In February 2001 funding of in total $1000k was received from DOE and a loan of $500k was raised from the U.S. CMS construction project. In April 2001 the remaining DOE funds of $500k were received. Also $320k were received from NSF another $80k were advised, to cover CAS engineering support. In the table those funds are shown as "Authorized". The budgeted funds shown are prorated for ongoing efforts.

All funds in AY$ x 1000	**FY2000**	**FY2001**		**FY2001Q1**	**FY2001Q2**
		Requested	Authorized	Budgeted	Budgeted
Core Application Software	870.0	1234.2	1234.2	268.7	288.6
User Facilities Personnel	260.7	885.4	635.4	158.9	158.9
User Facilities Hardware	344.0	1160.0	410.0	0.0	0.0
Project Office		120.4	120.4	0.0	0.0
Management Reserve		100.0	0.0	0.0	0.0
Total	**1474.7**	**3500.0**	**2400.0**	**427.5**	**447.4**

Table 3 *Funding status of the U.S. CMS Software and Computing Project in FY2001Q2. All funds are given in AY$ x 1000*

There was no news on the requested NSF funds for the Tier-2 hardware and personnel. The NSF support for CAS engineers was received. In Table 4 the funds received and the expected funds for FY2001 are summarized.

All funds in AY$ x 1000	**FY2000**	**FY2001**		**Total**
		Requested	Received	
DOE	1164.7	1500.0	1500.0	2664.7
NSF	310.0	1500.0	400.0	1810.0
Loan from U.S. CMS Detetector Project		500.0	500.0	500.0
Total	**1474.7**	**3500.0**	**2400.0**	**4974.7**

Table 4 *Funding sources, expected funds and funds received as of February 2001. All funds are given in AY$ x 1000*

IV. Milestone Status

WBS 1.4.1 Prototype Tier 2 Regional Centers

Milestone 1.4.1.2 Procurement Phase Complete. Completed 01/07/2001

Arrival of all components in first half of hardware deployment at Caltech and UC San Diego. Racks, Computational Nodes, Data Servers/Master Nodes, Network switches, RAID arrays. This milestone was delayed by one month due to lack of funds at UCSD.

Milestone 1.4.1.2.1 Commissioning and Testing of Half Size System. Completed 02/01/2001

Systems assembled, configured, cabled, and tested. CERN standard computing environment installed and CMS software compiled and installed.

Milestone 1.4.1.2.2 Reconstruction of events from CMS Fall Production: Due 03/21/01 Milestone missed

System was tested with CMSIM. GDMP was used to transfer minimum bias events for use as pile-up. Production scripts were deployed on Tier2 site. Full contribution to CMS fall production has not been realized due to lack of manpower (missing funding).

WBS 1.5.1 Onsite Networking Infrastructure

Milestone 1.5.1.1 CMS Tier1 Network Upgraded, Stage 1: Due 08/01/01

Design nearly complete. This will be a staged upgrade due to funding shortfalls

WBS 1.6.3 Computing and Software R&D: Fully Functional Regional Center

Milestone 1.6.3.1 Commissioning CMS Linux farms: Completed 11/01/2001

Milestone 1.6.3.1.1 Procurement of Test cluster: Due 07/01/2001
Milestone 1.6.3.1.2 Procurement of "User" cluster: Due 07/01/2001
Milestone 1.6.3.1.3 Procurement of additional CPU nodes: Due 07/01/2001

Bids for hardware are out.

WBS 1.7.1 Detector Construction Phase Computing: CMS Computing Facility

Milestone 1.7.1.2 Hosting User Federation at FNAL: Completed 02/20/2001
Adding events as they are reconstructed.

Milestone 1.7.1.3 Complete AOD: Due 01/30/2001
Overdue because of delays in fall production. Now in progress and we can project it will be finished around ~ 06/15/01

CAS Milestones:

CMS is in the process of overhauling the milestones and deliverables within the newly defined management structure of the Computing and Core Software (CCS) sub-project, as mentioned

above. This process will require modifications and consistency checks throughout the U.S. CMS CAS WBS. This is progressing but here is the milestone status for FY2001Q2 based on the last approved WBS presented at the baseline review.

WBS 2.1 Software Architecture

Milestone 2.1.2.1.2 Detector Description Conceptual Design: Due 03/11/2001

While progress is being made on the understanding of the underlying technologies, no consensus on a design and direction could be achieved. This milestone has slipped and the schedule is being modified to reflect the need for more study.

CMS is currently involved in a short investigational prototype of an XML detector description. The goal is to describe a test beam setup and access the information from OSCAR and ORCA for simulation and reconstruction and potentially use IGUANA for visualization. At the end of prototype a technological decision should be made about direction of the detector description. At this point U.S. CMS will assess whether to continue to direct our effort to this project.

Milestone 2.1.3.1.5 First Release of CAFÉ tools: Completed 03/01/2001
Released.

Milestone 2.1.3.2.1.4 Release of Top-Level Architectural Description Document for CARF: Due 03/01/2001
Work has progressed, but the document was not yet released.

WBS 2.2 Interactive Graphics and User Analysis

There are no major milestones in this item for FY2001Q2. The schedule for this task is being reworked to reflect the lack of available manpower, more tracking milestones will be added at this time.

WBS 2.3 Distributed Data Management and Processing

Milestone 2.3.1.1.4 Basic Functionality of Task Scheduling Prototype Deployed at Tier2 Prototype: Completed 02/01/2001
Milestone met.

Milestone 2.3.2.2.4 File Format Replication Deployed in GDMP: Due 02/01/2001

This has been delayed due to a delay in the release of the underlying technology, the Globus replica catalogue. There is a preliminary release that is being integrated into GDMP.

Milestone 2.3.2.4.4 Implementation of Security Protocol For Objectivity: Completed 04/01/2001

The implementation was available but unused until Objectivity version 6 was deployed for use in production. The spring CMS production relies on version 6.

Milestone 2.3.4.2.2 Generic Multi-site script development
First version released and adopted at several remote sites.